UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/655,564	09/03/2003	Adriana Dumitras	18602-08204	6437

61520 7590 03/31/2009

APPLE/FENWICK
SILICON VALLEY CENTER
801 CALIFORNIA STREET
MOUNTAIN VIEW, CA 94041

EXAMINER	
FINDLEY, CHRISTOPHER G	

ART UNIT	PAPER NUMBER
2621	

MAIL DATE	DELIVERY MODE
03/31/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

/Wanda L Walker/

Director, Technology Center 2600UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Application Number: 10/655,564
Filing Date: September 03, 2003
Appellant(s): DUMITRAS ET AL.

____

Brian G. Brannon (Reg. No. 57,219)
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 12/03/2008 appealing from the Office action mailed 5/02/2008.

(1) Real Party in Interest

Apple Inc. of Cupertino, California

(2) Related Appeals and Interferences

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) Status of Claims

The statement of the status of claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Claimed Subject Matter

The summary of claimed subject matter contained in the brief is correct.

(6) Grounds of Rejection to be Reviewed on Appeal

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

(7) Claims Appendix

The copy of the appealed claims contained in the Appendix to the brief is correct.

(8) Evidence Relied Upon

PCT/US97/08266	Chang et al.	11-1998
US 6,670,963 B2	Osberger	12-2003

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 101

1. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

2. **Claims 1-12 are rejected under 35 U.S.C. 101 as not falling within one of the four statutory categories of invention.**

Supreme Court precedent [Diamond v. Diehr, 450 U.S. 175, 184 (1981); Parker v. Flook, 437 U.S. 584, 588 n.9 (1978); Gottschalk v. Benson, 409 U.S. 63, 70 (1972); Cochrane v. Deener, 94 U.S. 780, 787-88 (1876)] and recent Federal Circuit decisions [In re Bilski, 88 USPQ2d 1385 (Fed. Cir. 2008)] indicate that a statutory "process" under 35 U.S.C. 101 must (1) be tied to another statutory category (such as a particular apparatus), or (2) transform underlying subject matter (such as an article or material) to a different state or thing. While the instant claim(s) recite a series of steps or acts to be performed, the claim(s) neither transform underlying subject matter nor positively tie to another statutory category that accomplishes the claimed method steps, and therefore do not qualify as a statutory process.

The method recited in claim 1 does not specify any physical components for carrying out each of the steps described. Furthermore, the method recited in claim 1 describes a process which may be carried out manually without the aid of a machine, as explained below.

A human being is capable of selecting two or more images from a set of multiple images when the sequence consists of still images, thereby satisfying the claim limitation of "selecting a set of frames from a video sequence."

Secondly, a human would be able to divide the images into uniform subsections and use his or her analytical capabilities to recognize subsections that are very similar between images. When the similar subsections are in different locations within their respective images, the human would then be able to manually create a vector by overlaying two sequential images and drawing arrows between the similar subsections. When this process is repeated for all of the subsections of each image, the human would have created a set of vectors for each image, thereby satisfying the claimed "determining a set of motion vectors for each frame in the set of frames."

Next, an angle for each of the drawn vectors may be measured by orienting all of the images, and thus the subsections with their corresponding vectors, according to Cartesian 2-dimensional x and y axes. The human would then be able to measure the angles of the vectors with respect to the axes by use of a protractor, thus satisfying the claimed "determining a motion angle for each motion vector."

The human would then be able to group the vectors into regions by designating each contiguous area of similarly oriented vectors as a region and determining the regions which contain the most subsections by counting the number of image subsections in each designated region, thereby satisfying the claimed "identifying at least two largest regions in each frame having motion vectors with substantially similar motion angles."

Next, the human would divide the number of image subsections in each of the previously determined largest designated regions by the total number of subsections in the image to identify the percentage of the image covered by each of the previously determined largest designated regions, thereby satisfying the claimed "determining percentages of each frame covered by the at least two largest regions;

The human would then be able to calculate an average angle of one of the previously determined largest designated regions by summing the angles of the constituent subsections for said designated region and dividing the sum by the number of subsections in the designated region. The human would next calculate the difference between the vector angle of each constituent subsection and the calculated average angle. The squared value of said difference may be calculated by the human, wherein the squared value of said difference is the variance for the designated region, thereby satisfying the claimed "determining a statistical measure of the motion angles for at least one of the two largest regions."

Finally, the sum of the calculated percentage for each of the previously determined largest designated regions may be compared to a threshold while the calculated variance value may be compared to a another threshold, as described in Equation 10 and 11 of the Applicant's specification, in order to determine the presence of pan or zoom in the image, thereby satisfying the claimed condition of "comparing the percentages and statistical measure to threshold values to identify at least one of a pan

and a zoom in the video sequence" and rendering claim 1 capable of being performed manually. Dependent claims 2-6 are also rejected accordingly.

The system recited in claim 7 does not specify any physical components for carrying out each of the operations described. Paragraphs [0025] and [0028] of the Applicant's specification indicate that the analysis performed by the system is implemented as software instructions stored on a computer readable medium and executed by one or more processors. However, the language presented in claim 7, and its dependent claims 8-12, simply references the analysis corresponding to software, which are not embodied in a computer readable medium.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 1-18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chang et al. (PCT/US97/08266) in view of Osberger (US 6670963 B2).

Re **claim 1**, Chang discloses a method of detecting at least one of a pan and a zoom in a video sequence, comprising: selecting a set of frames from a video sequence (In order to perform the analysis outlined in Chang it is inherent that a set of video frames are selected from a video sequence); determining a set of motion vectors for each frame in the set of frames (Chang: page 15, lines 10-12, each motion vector associated with the B and P frames contained in the shot are decoded); determining a motion angle for each motion vector (Chang: page 16, lines 3-6, the motion vectors of an object are represented in a 2-dimensional (x, y) plane, wherein it is inherent that each vector consists of a magnitude and direction, and the direction may be represented by an angle expressed as a geometric function of the x and y components); identifying at least two largest regions in each frame having motion

vectors with substantially similar motion angles (Chang: page 11, lines 10-15, The moving object(s) and the background constitute at least two regions, which are determined by their respective motion vector orientations; page 16, lines 3-4, each extracted object may be expressed by a global motion vector, indicating that the object is composed of several blocks with substantially similar motion angles/direction).

Chang does not explicitly disclose determining percentages of each frame covered by the at least two largest regions; determining a statistical measure of the motion angles for at least one of the two largest regions; and comparing the percentages and statistical measure to threshold values to identify at least one of a pan and a zoom in the video sequence. However, Osberger discloses a visual attention model, wherein motion vectors are calculated and used to determine pan, tilt, zoom, and rotate (Osberger: column 7, lines 23-28), and an estimate of the amount of motion in a scene is obtained by taking the mth percentile of the camera motion compensated motion vector map (Osberger: column 7, lines 61-64). Since both Chang and Osberger relate to determining pan and zoom in a video sequence, one of ordinary skill in the art at the time of the invention would have found it obvious to combine the statistical motion analysis of Osberger with the indexing and editing methods of Chang in order to provide an improved visual attention model of the type having spatial features to generate a spatial importance map and having a temporal importance map combined with the spatial importance map to produce an overall importance map for a frame (Osberger: column 8, lines 58-67), which would provide a more comprehensive technique for indexing video with key content browsing (Chang: page 6, lines 2-6). The combined method of Chang and Osberger has all of the features of claim 1.

Re **claim 2**, the combined method of Chang and Osberger discloses that the step of selecting a set of video frames from a video sequence further comprises: identifying a scene cut between two frames in the video sequence (Chang: Fig. 1, “SCENE CUT DETECTION” in PARSING section 110); and responsive to the identification of a scene cut, selecting a set of video frames from the video sequence that includes all the frames in the video sequence up to and including a frame just before the scene cut (Chang: column 2, lines 5-8).

Re **claim 3**, the combined method of Chang and Osberger discloses that frame differences and motion information are used to identify a scene cut (Chang: column 4, lines 45-61).

Re **claim 4**, the combined method of Chang and Osberger discloses that the motion angles are computed in one from the group of coordinate systems consisting of polar, Cartesian, spherical and cylindrical coordinate systems (Chang: Fig. 4 is a vector diagram, which serves to explain global and local motion. The x and y coordinates are contained in a Cartesian coordinate system and used in Equation (1) on page 14. Line 17 of page 14 states, "(x,y) is the coordinate of a macroblock in the current frame.").

Re **claim 5**, the combined method of Chang and Osberger discloses that the percentages of each flame covered by the at least two largest regions are determined from the number of pixels in each region as a percentage of the total number of pixels in a frame (Osberger: column 7, line 64, through column 8, line 9, pixel spacing and screen size are factored into the calculations).

Re **claim 6**, the combined method of Chang and Osberger discloses that the statistical measure is a variance (Osberger: column3, lines 50-55, variance is used in the calculations).

Claim 7 recites the corresponding system for implementing the method of claim 1, and, therefore, has been analyzed and rejected with respect to claim 1 above.

Claim 8 has been analyzed and rejected with respect to claim 2 above.

Claim 9 has been analyzed and rejected with respect to claim 3 above.

Claim 10 has been analyzed and rejected with respect to claim 4 above.

Claim 11 has been analyzed and rejected with respect to claim 5 above.

Claim 12 has been analyzed and rejected with respect to claim 6 above.

Claim 13 recites the corresponding computer readable medium stored thereon a computer program for executing the method of claim 1, and, therefore, has been analyzed and rejected with respect to claim 1 above.

Claim 14 has been analyzed and rejected with respect to claim 2 above.

Claim 15 has been analyzed and rejected with respect to claim 3 above.

Re **claim 16**, the combined method of Chang and Osberger discloses a majority of the features of claim 16, as discussed above regarding claims 1 and 13, but does not specifically disclose the use of polar coordinates in the motion vector analysis. However, the combined method of Chang and Osberger does use Cartesian coordinates (see fig. 4). The Examiner takes Official Notice that one of ordinary skill

in the art would have found it obvious to convert the Cartesian motion vector representation to polar coordinates as a personal preference for visual and/or mathematical representation. Polar coordinates provide no advantage over Cartesian coordinates and are simply a different way of representing the same data.

Claim 17 has been analyzed and rejected with respect to claim 5 above.

Claim 18 has been analyzed and rejected with respect to claim 6 above.

(10) Response to Argument

1. Regarding claims 1, 7, and 13, the Applicant contends (on page 6, lines 10-17, of the Applicant's appeal brief) that the prior art cited (Chang et al. (PCT/US97/08266) in view of Osberger (US 6670963 B2)) fails to teach or suggest the representative claim feature of "identifying at least two largest regions in each frame having substantially similar motion angles." In particular, the Applicant states that the detection disclosed by Chang "merely compares motion vectors to a predetermined threshold value to eliminate areas of the frame with motion vectors below the predetermined threshold value." However, the Examiner respectfully disagrees with this characterization.

Chang discloses extracting objects from a frame using global motion (Chang: page 16, lines 3-9), wherein the extraction involves separating foreground objects from the background (Chang: page 16, lines 9-20). Upon identification of foreground objects and background regions, Chang further discloses eliminating false small objects by comparing the motion vectors to a threshold to determine regions of little or no motion (Chang: page 17, lines 4-16). Chang also discloses nonlinear noise reduction by filtering out vectors that are very different from most of the other vectors in the same vicinity (Chang: page 15, lines 4-9), indicating that blocks with substantially similar vectors are grouped together.

If the frame consists of moving objects in the foreground with a non-moving background, the non-moving regions of a frame may be considered as being represented by a zero vector, wherein the vector has zero magnitude and no direction so that the lack of motion identifies non-moving blocks as belonging to a particular region of substantially similar zero motion vectors. Therefore, even in the instance where

one moving foreground object is extracted against a non-moving background, two regions are identified (the foreground object with its characteristic motion direction and magnitude and the background region with its characteristic zero vector magnitude and direction).

Since Chang indicates that substantially similar vectors are grouped together, as noted above, a frame with motion in the background would group the moving background as a region while any foreground objects would be extracted as separate regions as long as the foreground objects and background differ in their motion vectors. Each of said regions would be identified as having a unique motion parameter.

In either exemplary instance described, the background and foreground objects would be separately identified, meaning that at least two regions are identified. Since Change subsequently eliminates false small objects upon extraction of objects, only the largest objects are left once said small false objects are eliminated. Therefore, the conditions of "identifying at least two largest regions in each frame having substantially similar motion angles" are met by Chang.

2. Regarding claims 1, 7, and 13, the Applicant contends (on page 7, lines 3-11 of the Applicant’s appeal brief) that Osberger fails to remedy the alleged deficiency of Chang with respect to the limitation "identifying at least two largest regions in each frame having substantially similar motion angles." However, the Examiner respectfully submits that the alleged deficiency has already been addressed above.

3. Regarding claims 1, 7, and 13, the Applicant also contends (on page 7, line 3-page 10, line 3 of the Applicant’s appeal brief) that the prior art cited (Chang et al. (PCT/US97/08266) in view of Osberger (US 6670963 B2)) fails to teach or suggest the representative claim feature of "determining percentages of each frame covered by the at least two largest regions." In particular, the Applicant states, “Merely analyzing a specified percentile of a motion vector map does not determine of the percentages of a frame covered by the largest regions in the frame having motion vectors with substantially similar motion angles, but analyzes a subset of the motion vectors included in a complete frame, regardless of their location within the frame.” However, the Examiner respectfully disagrees.

Osberger discloses a visual attention model, wherein motion vectors are calculated and used to determine pan, tilt, zoom, and rotate (Osberger: column 7, lines 23-28), and an estimate of the amount of motion in a scene is obtained by taking the mth percentile of the camera motion compensated motion vector map (Osberger: column 7, lines 61-64). As discussed above in reference to Chang, the process of determining the at least two largest regions in each frame having substantially similar motion angles has already been addressed. Both Chang and Osberger disclose finding motion regions (Chang: page 16, lines 9-20 and Osberger: column 7, lines 23-28). Both Chang and Osberger also disclose performing a filtering operation for eliminating noisy areas (Chang: page 17, lines 4-16 and Osberger: column 7, lines 45-49). Osberger discloses an additional step of analyzing a motion vector map to determine a percentile (Osberger: column 7, lines 61-64). When the vector map containing the remaining motion vectors is analyzed, only the largest motion regions remain after the filtering. Accordingly, taking the percentile of motion in the frame identifies the amount of overall motion from the remaining largest regions. The Applicant contends that this overall motion for the remaining largest regions does not meet the claim limitations. However, the Applicant's own specification shows that the percentages of the largest regions are summed, and the sum is compared to a threshold (Applicant's specification: paragraph [0036] and equations 10 and 11), just as described in Osberger.

Also, Osberger states that adaptive thresholding is used to determine camera motion, wherein the threshold is lower for scenes with few and/or slow moving objects and the threshold is higher for scenes with many fast moving objects (Osberger: column 7, lines 55-61). The adaptive thresholds provide a more normalized look at the camera motion by compensating for the amount of object motion. Similarly, the Applicant's specification states that the threshold values can be adjusted (Applicant's specification: paragraph [0036]).

Since both Chang and Osberger relate to analyzing motion in video frames, one of ordinary skill in the art at the time of the invention would have found it obvious to combine the adaptive motion detection of Osberger with the scene content classification in order to provide a more robust way of detecting camera motion which is less susceptable to influence from high levels of motion variance

(Osberger: column 7, lines 13-21). Therefore, the Examiner respectfully submits that the combination of Chang and Osberger meets the limitations set forth in the claims.

4. The Applicant contends (on page 10, lines 15-22 of the Applicant's appeal brief) that the Official Notice relied upon by the Examiner with respect to claim 16 does not overcome the alleged deficiencies of Chang and Osberger. However, the Examiner respectfully submits that the alleged deficiencies have already been addressed above.

(11) Related Proceeding(s) Appendix

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

This examiner's answer contains a new ground of rejection set forth in section **(9)** above. Accordingly, appellant must within **TWO MONTHS** from the date of this answer exercise one of the following two options to avoid *sua sponte* **dismissal of the appeal** as to the claims subject to the new ground of rejection:

(1) **Reopen prosecution.** Request that prosecution be reopened before the primary examiner by filing a reply under 37 CFR 1.111 with or without amendment, affidavit or other evidence. Any amendment, affidavit or other evidence must be relevant to the new grounds of rejection. A request that complies with 37 CFR 41.39(b)(1) will be entered and considered. Any request that prosecution be reopened will be treated as a request to withdraw the appeal.

(2) **Maintain appeal.** Request that the appeal be maintained by filing a reply brief as set forth in 37 CFR 41.41. Such a reply brief must address each new ground of rejection as set forth in 37 CFR 41.37(c)(1)(vii) and should be in compliance with the other requirements of 37 CFR 41.37(c). If a reply brief filed pursuant to 37 CFR 41.39(b)(2) is accompanied by any amendment, affidavit or other evidence,

it shall be treated as a request that prosecution be reopened before the primary examiner under 37 CFR 41.39(b)(1).

Extensions of time under 37 CFR 1.136(a) are not applicable to the TWO MONTH time period set forth above. See 37 CFR 1.136(b) for extensions of time to reply for patent applications and 37 CFR 1.550(c) for extensions of time to reply for ex parte reexamination proceedings.

Respectfully submitted,

/Christopher Findley/

A Technology Center Director or designee must personally approve the new ground(s) of rejection set forth in section (9) above by signing below:

Conferees:

/Mehrdad Dastouri/

Supervisory Patent Examiner, Art Unit 2621

/Marsha D. Banks-Harold/

Supervisory Patent Examiner, Art Unit 2621

/Wanda L Walker/

Director, Technology Center 2600